REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	9/16/98	Final Progress Report

4. TITLE AND SUBTITLE
Crystal Structure of an Alcohol Dehydrogenase from the Extreme Thermophile *Thermoanaerobium brockii*

5. FUNDING NUMBERS
DAAH04-95-1-0396

6. AUTHOR(S)
Menachem Shoham

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Case Western Reserve University
Cleveland, OH 44106-4935

8. PERFORMING ORGANIZATION REPORT NUMBER
642-3043

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
ARO 34095.2-LS

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by the documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The crystal structure of a thermophilic alcohol dehydrogenase (TBAD) has been determined in complex with sec-butanol as substrate to a resolution of 3.0 Å. The enzyme is from *Thermoanaerobacter brockii*, a bacterium isolated from hot springs in Yellowstone National Park. This tetrameric enzyme has a melting temperature of 98 °C. The structure consists of two domains, the catalytic domain and the cofactor-binding domain, with the active site located in a deep cleft at the domain interface. Factors important for the thermostability have been deduced by comparison with the crystal structure of a highly homologous mesophilic alcohol dehydrogenase from *Clostridium beijerinckii* (CBAD). The thermophilic enzyme has a more hydrophilic exterior, a more hydrophobic interior, a smaller surface area, more prolines, alanines, and less serines than CBAD. Furthermore, in the thermophilic enzyme the number of all types of intersubunit interactions is increased: more salt bridges, hydrogen bonds and hydrophobic interactions. All these effects combined can account for the 32°C higher melting temperature of the thermophilic enzyme.

14. SUBJECT TERMS

15. NUMBER OF PAGES
27

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

TABLE OF CONTENS

RESEARCH BACKGROUND

The thermal stabilization of proteins is an intriguing scientific problem with important evolutionary and biotechnological implications (1-3). What mutation(s) should be engineered to increase the heat tolerance of a protein? There may be no unique answer to this question. Structural studies of thermophilic proteins and their mesophilic counterparts have indicated many small contributions but no single structural feature that could account for the increased thermostability (4-9). A powerful way to stabilize oligomeric proteins is to increase the number of intersubunit interactions. However, in a hyperthermophilic organism all proteins have to be heat resistant, even monomeric ones. Factors known to contribute to thermal stability include more hydrophobic contacts in solvent-inaccessible areas, increased number of salt bridges, more hydrogen bonds, more proline residues and a higher degree of compactness.

In an attempt to broaden our understanding of the principles of protein adaptation to high temperatures, we have undertaken the crystal structure determination of an alcohol dehydrogenase (TBAD) from *Thermoanaerobacter brockii*, a bacterium isolated from hot springs in Yellowstone National Park. TBAD is an NADP (nicotinamide adenine dinucleotide phosphate) dependent enzyme, which reversibly catalyzes the oxidation-reduction reactions of alcohols to their corresponding aldehydes or ketones (10). The best substrates are secondary alcohols. The temperature at which half the activity is lost after one hour of incubation is 93°C, and the melting temperature is 98°C (11). Another unusual property of this extremophilic enzyme is its high tolerance towards organic solvents. TBAD has been used in organic synthesis as a biocatalyst for the stereospecific reduction of a broad range of ketones to their corresponding secondary alcohols (12, 13). TBAD is a tetrameric enzyme of identical 37,652 Da subunits composed of 352 amino acids (14). In contrast to eukaryotic alcohol dehydrogenases, which contain a catalytic Zn^{2+} as well as a structural Zn^{2+}, TBAD has only a single Zn^{2+} ion per subunit, reported to be of catalytic

function (15). The ligands to the Zn^{2+} have been identified by alanine mutagenesis as Cys37, His59, and Asp150 (15).

The crystal structures of several eukaryotic alcohol dehydrogenases have been determined, including those from cod and horse liver (CLAD and HLAD) (16, 17), as well as three human recombinant forms (18-20). The ligands to the catalytic Zn^{2+} in eukaryotic alcohol dehydrogenases are a histidine residue and two cysteine residues. The fourth ligand is usually a water molecule in the absence of substrate. When the substrate is bound to the enzyme, the alcohol oxygen or carbonyl oxygen atom appears to coordinate directly to the Zn^{2+}, displacing the molecule of water. Studies of the pH dependence of binding of a variety of alcohols to HLAD indicate that the alcohol probably binds in the form of alcoholate anion (21). The negative charge on the oxygen atom makes the alcoholate a better donor of a hydride ion for the reduction of NAD^+, compared with the undissociated alcohol. The positively charged zinc ion stabilizes the negative charge of the alcoholate in HLAD.

The sequence of TBAD is remotely related to those of eukaryotic alcohol dehydrogenases with a degree of sequence identity around 30%. In contrast, TBAD is highly homologous to a recently crystallized prokaryotic alcohol dehydrogenase from Clostridium beijerinckii (CBAD) with a sequence identity of 75% (22). CBAD and TBAD differ in their thermal stability and in their resistance to denaturing agents. The extent of resistance to denaturation in TBAD depends on the presence of the cofactor (22). The temperature at which half the activity is lost within one hour is 67 °C and 93 °C for CBAD and TBAD, respectively (22). The increased thermostability in TBAD must be localized to those 25% of the residues, which are different in the two enzymes.

We report here the crystal structure of TBAD in a binary substrate complex with sec-butanol to a resolution of 3.0 Å. Catalysis does not take place in these crystals due to the absence of cofactor.

EXPERIMENTAL PROCEDURES

Crystallization, X-ray data collection and processing

TBAD was obtained from the laboratory of Dr. Yigal Burstein at the Weizmann Institute of Science. TBAD crystals were grown by the vapor diffusion method against a well solution containing 10% PEG-8000 (polyethylene glycol), 5.8% (v/v) MPD (methyl pentane diol), 2.5% (v/v) sec-butanol, 50 mM magnesium acetate, 0.1 M MES (2[N-Morpholino]ethanesulfonic acid) buffer, pH 5.8. The crystals belong to space group $P2_12_12_1$, with unit-cell dimensions a=80.45 Å, b=123.08 Å, c=168.03 Å. There are four molecules of TBAD per asymmetric unit. The volume-to-mass ratio is 2.86 $Å^3Da^{-1}$, which is close to the average for protein crystals (23). The solvent content is 57%.

Prior to data collection, the crystals were dialyzed overnight against a cryoprotectant solution containing 10% PEG-8000, 20% (v/v) MPD, 2.5% sec-butanol, 50 mM magnesium acetate, 0.1 M MES buffer, pH 5.8. X-ray diffraction data were collected at CHESS beamline A1 using synchrotron radiation (λ=0.908 Å) at 100 °K, with a crystal-to-detector distance of 253 mm. The crystals diffracted to 2.99 Å. X-ray diffraction data were processed using program DENZO and reduced with SCALEPACK (24). Table 1 shows the statistics of data collection and processing.

Structure determination and refinement

The structure of TBAD was determined by molecular replacement with the program AmoRe (25) within the CCP4 program package (26), using data in the resolution range from 15 to 4 Å. Reflections with $F>\sigma_F$ were used throughout the molecular replacement and subsequent refinement calculations. The crystal structure of the CBAD monomer (PDB code 1KEV) devoid of NADP and water molecules was chosen as a starting model (27). The rotation search solutions were somewhat noisy. Translation-function searches were subsequently carried out at the same resolution range for

the top 35 rotation solutions. Four of the solutions, ranked 1, 3, 13, and 26 in the rotation search, were found to be outstanding in the translation search with correlation coefficients at least 38% higher than the rest of solutions. The correctness of these four solutions was validated by the fact that they pack well to form the tetramer, and they conform to the subunit relationships revealed by a self-rotation search (data not shown). The tetramer has non-crystallographic 222 symmetry. The model of the tetramer generated from the four top solutions yielded an R-factor of 36.1% and a correlation coefficient of 63.6%.

The molecular-replacement model was subsequently modified by mutating non-conserved residues to Ala and setting all B-factors to 20 $Å^2$. Refinement was carried out using X-PLOR, versions 3.1 and 3.8 (28). Rigid-body refinement was carried out to improve the positional and orientational parameters of the four subunits. Strict non-crystallographic symmetry was maintained in this and all subsequent refinement cycles. A series of twenty-residue omit (F_o-F_c) maps were calculated and used to retrace the whole model using the TBAD amino acid sequence. The Zn^{2+} ion and the substrate were located from a residual (F_o-F_c) map after all protein residues were accounted for. Manual rebuilding between refinement cycles was performed with the computer-graphics software O (29). The model was subjected to several cycles of simulated annealing by "heating" the system to 3500 °K and slowly "cooling" it to 300 °K in time steps of 0.00025 ps. This reduced the R-value to 26.3%. Positional parameters and temperature-factors were then refined in alternative cycles. A ($2F_o$-F_c) Fourier map was computed at every stage of the refinement and the model was adjusted manually to fit this electron-density map. Water molecules were not added, except for the one in close proximity to the zinc ion. Side chain density was good for all residues except for a small portion of side chains on the surface. The Ramachandran plot shows that 84.6% of the residues are in most favored regions with none in the disallowed regions, as calculated with

program PROCHECK (30). The statistics of the crystallographic analysis and the final model are listed in Table 2. The coordinates will be deposited at the Protein Data Bank.

SUMMARY OF THE MOST IMPORTANT RESULTS

Structure description

As in all other alcohol dehydrogenases of known structure, the monomer is made up of two domains, the catalytic domain and the cofactor-binding domain (Fig. 1). The catalytic domain contains the N- and C-terminal segments, consisting of residues 1 to 149 and 297 to 352. There are two β-sheets and five α-helices in the catalytic domain. Strands 1 and 2 form a small antiparallel β-meander on the surface of this domain, while strands 3,4,5,12 and 13 form a mixed β–sheet. The loop between strands 3 and 4 is very long and contains a helical stretch α1. Following β5 there are four α-helical segments in a row. The last of these, α5, links the catalytic domain to the nucleotide-binding domain.

The nucleotide-binding domain consists of the contiguous stretch of residues 157 to 292. It forms a typical Rossmann fold with six parallel β-strands, β6 to β11. This β-sheet is flanked by α-helical segments, α6 and α7 on one side of the sheet, and α8 and α9 on the other side of the sheet. Following the last strand in this sheet, β11, there is a long loop which crosses over back into the catalytic domain to lead into helix α10, strand β12, helix α11 and finally strand β13. The two domains are separated by a deep active site cleft that is accessible to the solvent.

The active site

The location of sec-butanol and the zinc ion in the active site cleft was determined from a residual (F_o-F_c) difference Fourier map (Fig. 2). The binding site for the substrate sec-butanol consists of a hydrophobic region of residues Ile86, Leu107, Tyr267, and Leu294, as well as the

hydrophilic residues His59 and Asp150 (Fig. 3). The hydroxyl group of the substrate is hydrogen-bonded to Nε2 of His59 and one of the carboxylate oxygen atoms of Asp150 with distances of 3.2 and 3.0 Å, respectively. These residues have been shown to be essential for catalysis (15). Surprisingly, the zinc ion is not in direct proximity to the substrate but it is located some 9 Å away from the hydroxyl oxygen atom of the substrate. Residues surrounding the zinc ion include Cys37, Thr38, Met337, Met151 and a water molecule (Table 3). His59 and Asp150 are not in direct contact with the zinc ion. However, if the zinc ion were to be involved in catalysis, a conformational change would be required, presumably upon NADP binding, which would bring the zinc ion closer to the substrate. Indeed, substantial conformational changes between the apo and holo forms of liver alcohol dehydrogenases have been observed (31).

Structural features that account for the thermostability

The degree of sequence identity between TBAD and CBAD is 75%. Factors that account for the increased thermostability in TBAD must therefore be located in the 25% of the residues that are different between the two enzymes (Fig. 4). The sequence alignment is unambiguous since there are no gaps. There are a total of 87 residues that are different between the two enzymes. As shown in Table 4, the most significant differences in amino acid composition between the two enzymes are for alanine, proline and serine.

Proline residues—The number of proline residues increases from 13 in the mesophilic enzyme to 21 in the thermophilic enzyme (Table 4). Proline reduces the main-chain flexibility of the backbone, thereby decreasing the entropy and increasing the free energy of the unfolded state (32). Consequently any replacement to a proline should destabilize the unfolded state, thereby stabilizing the folded state. Interestingly, these replacements do not occur at random positions but at preferential sites. Of the eight extra prolines in TBAD, Pro177, Pro222 and Pro316 are found at the

N-terminal position of helices. A residue in this position is partly outside the helix and partly inside, i.e. the proline ring is not part of the helix but the carbonyl oxygen forms a α-helical hydrogen bond with the amide nitrogen four residues upstream. Proline residues rarely occur in helix interiors, where they would cause a kink in the helix (33, 34). Side chain hydrogen bonding to the backbone in a helix is rarely to the amide nitrogen of the first residue but frequently to the second or third residue. The lack of a hydrogen atom on the amide nitrogen of a proline residue in position one of a helix seems therefore not be a hindrance. Single proline site directed mutagenesis of CBAD revealed the following changes in the melting temperature: an increase for the alanine-to-proline substitution at position 177 and the leucine-to-proline substitution at position 316 of 0.5 °C and 10.8 °C, respectively; and a decrease for the histidine-to-proline substitution at position 222 of 6.5 °C (11). Position 222 is on the surface of the protein. Replacing a histidine for a proline at this position leads to a loss of solvation energy, which would account for the observed destabilization. Pro316 is the first residue in a 4-residue long stretch of 3_{10} helical conformation, linking an α–helix to a β-strand. Replacing a bulky leucine for the smaller proline may relieve some overcrowding in this region containing Met147, Met152 and Val309. Therefore, one may conclude that the introduction of a proline at the beginning of a helix does not stabilize the protein unless it is a special position. Pro24 is located at position two in a β-turn (Fig. 5), a position that is frequently populated by prolines. The serine-to-proline mutant of CBAD at this position has a 3.9°C higher melting temperature (11). The introduction of a kink at this position is facilitated by the fixed Φ angle in proline.

Alanine and serine residues—As shown in Table 4, TBAD has eight more Ala residues compared to CBAD. Three of the replacements to alanine occur in α-helical regions. An alanine increases the stability of a helix because of the unrestricted rotational freedom of the methyl side chain (35, 36). Introduction of alanine residues at other positions involve replacement of polar side

chains such as serine and arginine in buried regions (Ser35, Ser168, Ser246, Ser250, Ser263, Arg238), as well as replacement of valine with the smaller and less hydrophobic alanine on the surface of the protein at position 32 (Fig. 4). These replacements also provide entropic stabilization. Some serine residues in the core of the protein are replaced with the more apolar threonine residues without loss of hydrogen bond capacity (Ser92, Ser154, and Ser169). Two exposed serine residues at positions 254 and 270 are replaced with the more hydrophilic lysine and glutamic acid, respectively.

Surface characteristics and subunit interfaces—TBAD is more compact than its mesophilic counterpart. The accessible surface area in TBAD is 7.7% smaller than in CBAD (Table 5). Tighter packing has indeed been suggested as a mechanism for enhanced thermostability in proteins (4). Moreover, the area buried upon oligomerization is significantly larger in TBAD (Table 5); i.e. there are more extensive contacts between the subunits in the thermophilic enzyme. These intermolecular contacts include salt bridges, hydrogen bonds and hydrophobic interactions.

Salt bridges, hydrogen bonds and van der Waals contacts—A major cause of higher thermostability can be attributed to an increase in the number of van der Waals contacts, hydrogen bonds and salt bridges, particularly those at subunit interfaces (Table 6). Although the TBAD monomer has fewer hydrogen bonds in comparison to CBAD, due to the loss of buried polar residues, this is more than offset by the creation of extra salt bridges and van der Waals contacts both within and between the monomers. Furthermore, additional intermolecular hydrogen bonds are gained in TBAD. Salt bridges are important stabilizing forces in proteins if they occur in regions inaccessible to the solvent (37-39). The strength of a salt bridge depends on the environment, i.e. the local dielectric constant, and the bond distance. There are 31 salt bridges in TBAD versus 24 in CBAD (Table 6). Some of the salt bridges in TBAD are involved in extensive networks. The energy gain in a network is larger than the sum of the pairwise interactions between the constituent

charges (40). An example of a sequestered network of six salt bridges - some intramolecular and some intermolecular – is depicted in Fig. 6. Only two of these salt bridges exist in CBAD. Hydrogen bonds occur between a hydrogen bond donor and acceptor pair with a distance of less than 3.3 Å at an angle greater than 90 °(4). There are 218 intermolecular hydrogen bonds in TBAD versus 167 in CBAD (Table 6). Accurate accounting of salt bridges and hydrogen bonds in TBAD is difficult, as the crystal structure was determined at 3.0 Å. A few genuine interactions may have been missed, and the validity of other assigned bonds may be questioned due to large interatomic distances. However, significant differences in the number of interactions between the thermophilic and the mesophilic enzyme clearly suggest mechanisms to enhance the thermal stability of TBAD.

BIBLIOGRAPHY

1. Goodenough, P. W. (1995) Mol. Biotechnol. 4, 151-166
2. Lasa, I., and Berenguer, J. (1993) Microbiologia 9, 77-89
3. Coolbear, T., Daniel, R. M., and Morgan, H. W. (1992) Adv. Biochem. Eng. Biotechnol. 45, 57-98
4. Wallon, G., Kryger, G., Lovett, S., Oshima, T., Ringe, D., and Petsko, G. A. (1997) J. Mol. Biol. 266, 1016-1031
5. Matthews, B. W. (1995) Adv. Protein Chem. 46, 249-278
6. Rees, D. C., and Adams, M. W. (1995) Structure 3, 251-254
7. Spassov, V. Z., Karshikoff, A. D., and Ladenstein, R. (1995) Protein Sci. 4, 1516-1527
8. Day, M. W., Hsu, B. T., Joshua-Tor, L., Park, J. B., Zhou, Z. H., Adams, M. W., and Rees, D. C. (1992) Protein Sci. 1, 1494-1507
9. Russell, R. J., and Taylor, G. L. (1995) Curr. Opin. Biotechnol. 6, 370-374
10. Lamed, R. J., and Zeikus, J. G. (1981) Biochem. J. 195, 183-190

11. Bogin, O, Peretz, M., Hacham, Y., Korkhin, Y., Frolow, F., Kalb, A. J., and Burstein, Y. (1998) Protein Sci. 7, 1156-1163

12. Keinan, E., Seth, K. K., Lamed, R., Ghirlando, R., and Singh, S. P. (1990) Biocatalysis 3, 57-71

13. Keinan, E., Sinha, S. C., and Sinha-Bagchi, A. (1991) J. Chem. Soc. Perkin Trans. 1, 3333-3339

14. Peretz, M., and Burstein, Y. (1989) Biochemistry 28, 6549-6555

15. Bogin, O., Peretz, M., and Burstein, Y. (1997) Protein Sci. 6, 450-458

16. Ramaswamy, S., Ahmad, M. E., Danielsson, O., Jornvall, H., and Eklund, H. (1996) Protein Sci. 5, 663-671

17. Eklund, H., Ramaswamy, S., Plapp, B. V., el-Ahmad, M., Danielsson, O., Hoog, J. O., and Jornvall, H. (1994) EXS 71, 269-277

18. Davis, G. J., Bosron, W. F., Stone, C. L., Owusu-Dekyi, K., and Hurley ,T. D. (1996) J. Biol. Chem. 271, 17057-17061

19. Xie, P., Parsons, S. H., Speckhard, D. C., Bosron, W. F., and Hurley, T. D. (1997) J. Biol. Chem. 272, 18558-18563

20. Hurley, T.D., Bosron, W.F., Stone, C.L., and Amzel, L.M. (1994) J. Biol. Chem. 23, 415-429

21. Eklund, H., Plapp, B. V., Samama, J. P., and Branden, C. I. (1982) J. Biol. Chem. 257, 14349-14358

22. Korkhin, Y., Frolow, F., Bogin, O., Peretz, M., Kalb, A. J., and Burstein, Y. (1996) Acta Cryst. D52, 882-886

23. Matthews, B.W. (1968) J. Mol. Biol. 33,491-497

24. Otwinowski, Z. (1993) in Proceedings of the CCP4 Study Weekend: Data Collection and Processing (L Sawyer, N. Isaacs & S. Bailey eds) pp.56-62, SERC Daresbury Laboratory, Warrington, UK

25. Navaza, J. (1994) Acta Crystallog. A50, 157-163

26. Collaborative Computational Project, Number 4 (1994) Acta Crystallog. D50, 760-763.

27. Korkhin, Y., Kalb (Gilboa), A. J., Peretz, M., Bogin, O., Burstein, Y., and Frolow, F. (1998) J. Mol. Biol. 278, 967-981

28. Brunger, A. T. (1993) X-PLOR, Version 3.1 manual: a system for X-ray crystallography and NMR, Yale University Press, New Haven, USA

29. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallog. A47, 110-119

30. Laskowski, R. A., MacArthur, M. W., Moss, D. S , and Thornton, J. M. (1993) J. Appl. Crystallog. 26, 283-291

31. Collona-Cesari, F., Perahia, D., Karplus, M., Eklund, H., Branden, C. I., and Tapia, O. (1986) J. Biol. Chem. 261, 15273-15280

32. Matthews, B. W., Nicholson, H., and Becktel W. J. (1987) Proc. Natl. Acad. Sci. 84, 6663-6667

33. Richardson, J. S., and Richardson, D. C. (1988) Science 240,1648-1652

34. Aurora, R., and Rose, G. D. (1998) Protein Sci. 7, 21-38

35. Menendez-Arias, L., and Argos, P. (1989) J. Mol. Biol. 206, 397-406

36. Kelly, C. A., Nishiyama, M., Ohnishi, Y., Beppu, T., and Birktoft J. J. (1993) Biochemistry 32, 3913-3922

37. Nakamura, H. (1996) Q. Rev. Biophys. 29, 1-90

38. Shortle, D. (1992) Q. Rev. Biophys. 25, 205-250

39. Barlow, D. J., and Thorton, J. M. (1983) J. Mol. Biol. 168, 867-885

40. Horovitz, A., Serrano, L., Avron, B., Bycroft, M., and Fersht, A. R. (1990) J. Mol. Biol. 216, 1031-1044

41. Kraulis, P. J. (1991) J. Appl. Crystallog. 24, 946-950

42. Merritt E. A., and Murphy M. E. (1994) Acta Crystallog. D50, 869-873

TABLE I. Statistics of Data Collection and Processing

Resolution (Å)	Reflections	Completeness		redundancy		<I/σ>		R_{merge}*
		all	3.10-2.99	all	3.10-2.99	all	3.10-2.99	
2.99	30323	0.863	0.890	4.4	4.8	12.1	4.6	0.083

*$R_{merge} = \Sigma_h\Sigma_i|I_{ih}-\langle I_h\rangle|/\Sigma_h\Sigma_i\langle I_h\rangle$ where $\langle I_h\rangle$ is the mean intensity of the i observations of reflection h.

TABLE II Refinement Statistics

Resolution (Å)	50-2.99
Reflections (work/free)	26756/2986
R^*/R_{free}**	0.211/0.265
No. of protein atoms	2638
r.m.s. bond lengths (Å)	0.02
r.m.s. bond angles (°)	1.89
Avg. B-factors (main/side) ($Å^2$)	19.53/23.15

* $R = \dfrac{\Sigma|F_o - F_c|}{\Sigma F_o}$ F_o, observed structure factor; F_c calculated structure factor

** R_{free}, defined as in R for a randomly selected set of 10% of the reflections

TABLE III Distances between the zinc ion and surrounding residues (Å)

Cys37 Sγ	3.2	Thr38 Oγ1	4.4
Met337 Sδ	3.6	Water O	3.0
Met151 Sδ	4.9		

TABLE IV Differences in the amino acid composition

	TBAD	CBAD	TBAD-CBAD
Ala	35 (.099)	27 (.077)	8 (0.022)
Arg	14 (.040)	14 (.040)	0
Asn	10 (.028)	13 (.037)	-3 (-0.009)
Asp	20 (.057)	21 (0.060)	-1 (-0.003)
Cys	4 (0.011)	5 (0.014)	-1 (-0.003)
Gln	3 (0.009)	6 (0.017)	-3 (-0.008)
Glu	21 (0.060)	18 (0.051)	3 (0.009)
Gly	43 (0.122)	45 (0.128)	-2 (-0.006)
His	10 (0.028)	11 (0.03.1)	-1 (-0.003)
Ile	26 (0.074)	26 (0.074)	0
Leu	23 (0.065)	28 (0.080)	-5 (-0.015)
Lys	24 (0.068)	21 (0.060)	3 (0.008)
Met	15 (0.043)	18 (0.051)	-3 (-0.008)
Phe	14 (0.040)	11 (0.031)	3 (0.009)
Pro	21 (0.060)	13 (0.037)	8 (0.023)
Ser	9 (0.026)	17 (0.048)	-8 (-0.022)
Thr	13 (0.037)	11 (0.031)	2 (0.008)
Trp	4 (0.011)	4 (0.011)	0
Tyr	6 (0.017)	7 (0.020)	-1 (-0.003)
Val	37 (0.105)	35 (0.100)	2 (0.005)

The amino acid composition is based on SWISSPROT entries P14941 and PP25984 for TBAD and CBAD, respectively. Values in parentheses represent mol fractions. Differences larger than one percent are shown in bold.

TABLE V Surface area of the tetramer

	TBAD	CBAD	Difference (%)
Total accessible surface area (Å^2)	44996	48775	-7.7
Area buried in tetramer interfaces (Å^2)	14684	14397	2.0
Fraction of buried area	0.326	0.295	3.1
Percentage of exposed hydrophilic residues	71.4	69.2	2.2
Percentage of exposed hydrophobic residues	28.6	30.8	-2.2
Percentage of buried hydrophilic residues	19.4	24.5	-5.1
Percentage of buried hydrophobic residues	80.6	75.5	5.1

TABLE VI Interactions in CBAD and TBAD

	Intermolecular contacts			Intramolecular contacts		
	H-bonds	salt bridges	van der Waals	H-bonds	salt bridges	van der Waals
CBAD	167	8	774	538	16	4383
TBAD	218	13	862	476	18	4689
Difference	51	5	88	-62	2	306

The contacts were calculated with program CONTACT within CCP4 crystallographic suite of programs (26). Threshold values: salt bridges 4.0, H-bonds 3.4, van der Waals contacts 4.0 Å.

FIG. 1. Ribbon drawing of the TBAD monomer. Helices and strands are shown in dark and light gray, respectively. The enzyme is made up of two domains, the nucleotide-binding domain and the catalytic domain. The two domains are separated by a deep active site cleft that is freely accessible to the solvent. The Zn^{2+} ion is located at the bottom of the cleft in the catalytic domain. This figure was created with programs MOLSCRIPT (41) and RASTER3D (42).

FIG. 2. Residual ($F_o - F_c$) difference Fourier map contoured at a level of four standard deviations above the mean. F_c contains contributions from all protein atoms but no other atoms. Peak1 has been interpreted as the zinc ion due to its proximity to Cys37. Peak2 has been interpreted as sec-butanol due to its proximity to the catalytic residues His59 and Asp150 and a hydrophobic region of residues Ile86, Leu107, and Leu294. These are the two most prominent peaks in this map. Residues in the vicinity of the active site are depicted in ball-and-stick. This drawing was made with program O (29).

FIG. 3. Ball-and-stick representation of residues in the vicinity of the active site. The binding site for the substrate sec-butanol consists of a hydrophobic region including residues Ile86, Leu107, Tyr267 (not shown for clarity) and Leu294, as well as the catalytic residues His59 and Asp 150. The zinc ion is not in direct proximity to the substrate but it is located some 9 Å away from the hydroxyl oxygen atom of the substrate. Residues surrounding the zinc ion include Cys37, Thr38, Met151, Met337, and a water molecule.

FIG. 4. Sequence Alignment of TBAD with CBAD. The lines marked "TBAD" contain only changes from CBAD. Residues marked "b" are buried (accessibility < 25 $Å^2$), residues marked "e" are exposed in the tetramer (accessibility > 25 $Å^2$).

FIG. 5. Ball-and-stick representation of a region containing three proline residues in the thermophilic enzyme at positions 20, 22 and 24. TBAD is shown in red and CBAD in cyan. Pro20 is conserved in both enzymes whereas residues 22 and 24 in CBAD are alanine and serine, respectively. Introduction of a proline in position 22 increases the number of van der Waals contacts with nearby Ile28. Residue 24 is located at position two of a β-turn, a position that is frequently populated by prolines. There is little or no entropy loss during folding for a proline at this position.

FIG. 6. Ball-and–stick representation of the extensive network of salt bridges involving three subunits in TBAD. Residues are labeled red, green, and white, according to the subunit they belong to. R97 from the "white" subunit forms an intramolecular salt bridge with E94 (shortest distance 3.4Å). Across the subunit interface from R97 the carboxylate of D237 from the "red" subunit is located at a distance of 5.8 Å, normally considered too long for a salt bridge. However, in a milieu of low dielectric constant these two opposite charges might still exert some attractive force on each other. This same residue D237 also interacts with K257 from the same subunit as well as with R304 from the "green" subunit at distances of 3.1 and 3.4 Å, respectively. R304 from the "green" subunit in turn interacts with E165 from the "red" subunit at a distance of 4.1 Å. The carboxylate of E165 from the "red" subunit is 7.1 Å from the guanido group of R301 from the "green" subunit, too far to be considered a salt bridge. Finally, this R301 interacts with E160 from the same "green"

subunit at a shortest distance of 2.5 Å. There are five salt bridges depicted in this figure – six if you count the long intermolecular bridge between R97 and D237. Of these only two are present in mesophilic CBAD.

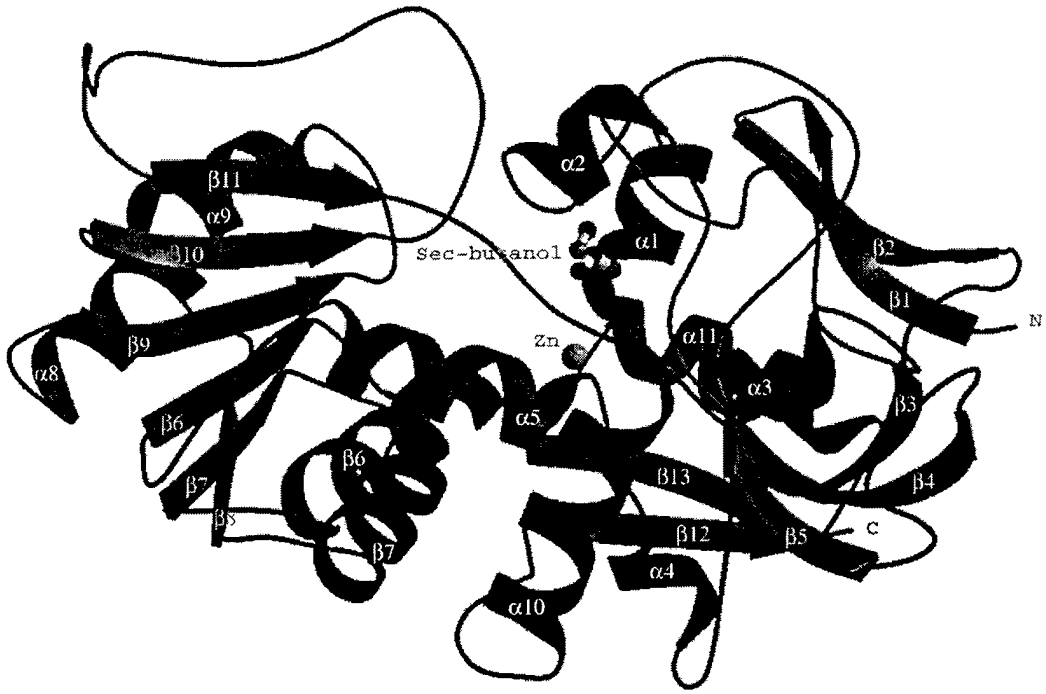
β11
α9
β10
β9
α8
β6
β7
β8
β6
β7
α2
α1
Sec-butanol
Zn
α11
α3
α5
β13
β12
α4
α10
β2
β1
N
β3
β4
β5
C

Fig. 1

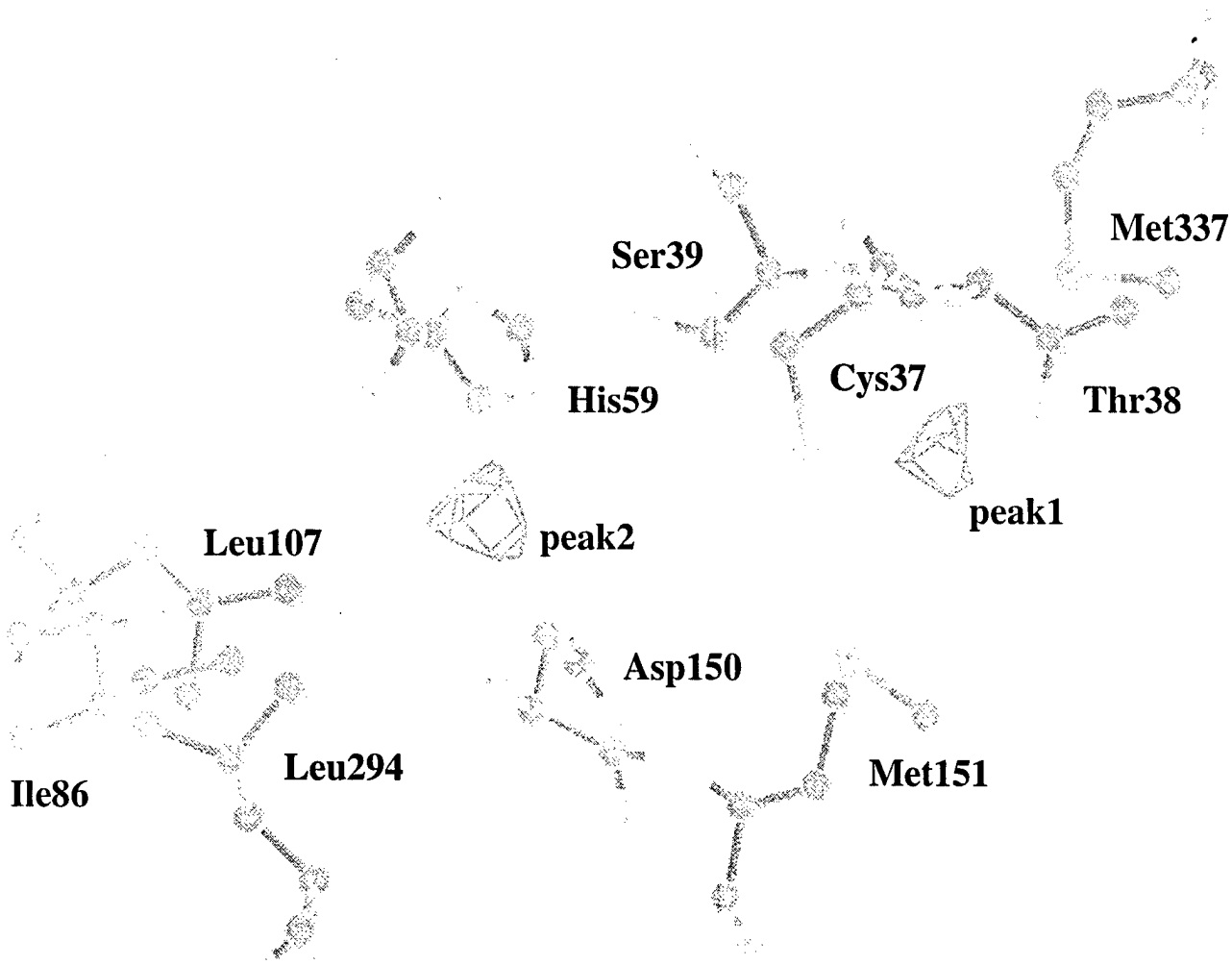
Met337
Ser39
His59
Cys37
Thr38
peak1
peak2
Leu107
Asp150
Leu294
Met151
Ile86

Fig. 2

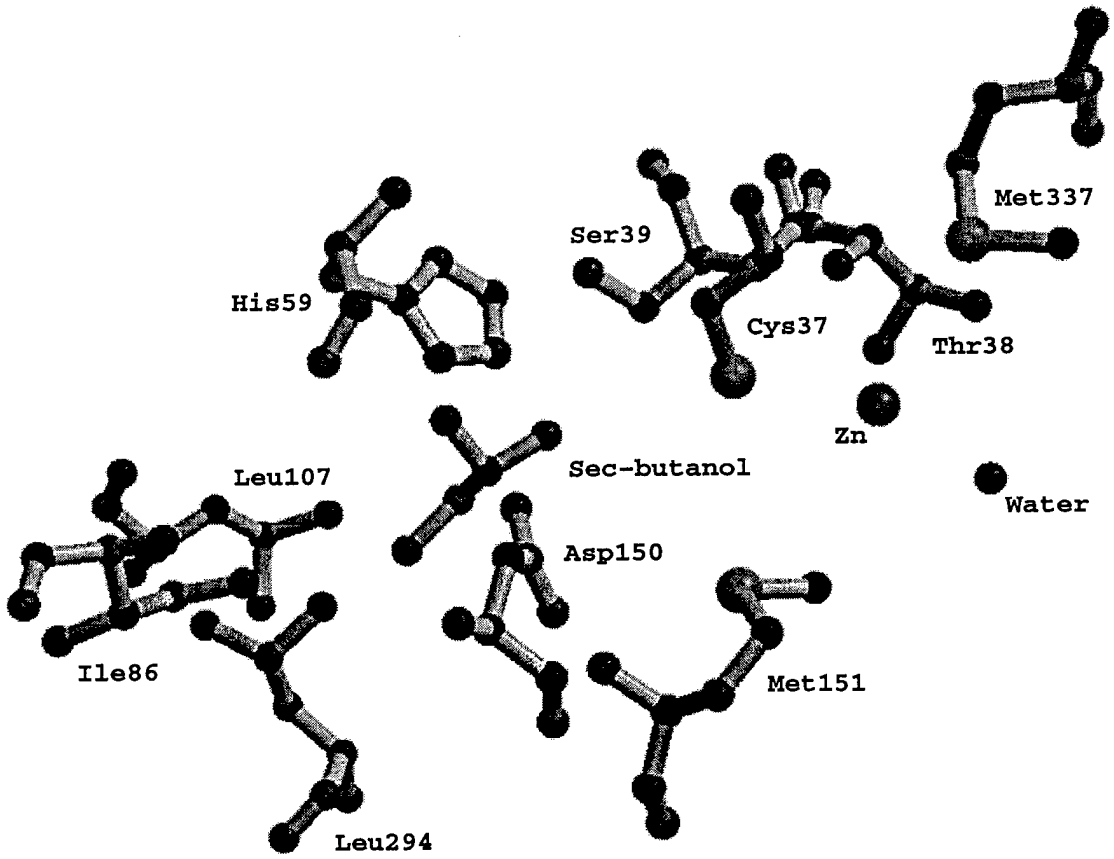
Met337
Ser39
His59
Cys37
Thr38
Zn
Sec-butanol
Water
Leu107
Asp150
Ile86
Met151
Leu294

Fig. 3

```
CBAD      1 MKGFAMLGINKLGWIEKERPVAGSYDAIVRPLAVSPCTSDIHTVFEGALGDRKNMILGHE
TBAD      1         S G V      K AP PF         A              I E H
                    e e b      e ee ee         b              e e e

CBAD     61 AVGEVVEVGSEVKDFKPGDRVIVPCTTPDWRSLEVQAGFQQHSNGMLAGWKFSNFKDGVF
TBAD     61                          V  AI     TS   R YH   G          V
                                     b  bb     be   e ee   e          e

CBAD    121 GEYFHVNDADMNLAILPKDMPLEAAVMITDMMTSGFHGAELADIQMGSSVVVIGIGAVGL
TBAD    121   F           H   EI        P    T          EL AT A L   P
                e           e   eb   b    b    b          eb bb b b   b

CBAD    181 MGIAGAKLRGAGRIIGVGSRPICVEAAKFYGATDILNYKNGHIVDQVMKLTNGEGVDRVI
TBAD    181  AV            A     V  D   Y      V   D P ES I N  E K   AA
             bb            b     e  e   e      b   e e ee b e  e e   bb

CBAD    241 MAGGGSETLSQAVSMVKPGGIISNINYHGSGDALLIPRVEWGCGMAHKTIKGGLCPGGRL
TBAD    241 I   NADIMAT  KI     T A V  F E EV PV  L
             b   ebebbbb  eb     b b b  b e ee eb  e

CBAD    301 RAEMLRDMVVYNRVDLSKLVTHVYHGFDHIEEALLLMKDKPKDLIKAVVIL-
TBAD    301  M R I L F K   P       FR   N  K FM          P    A
             b e b b e e   b       ee   e  e be          b    e
```

Fig. 4

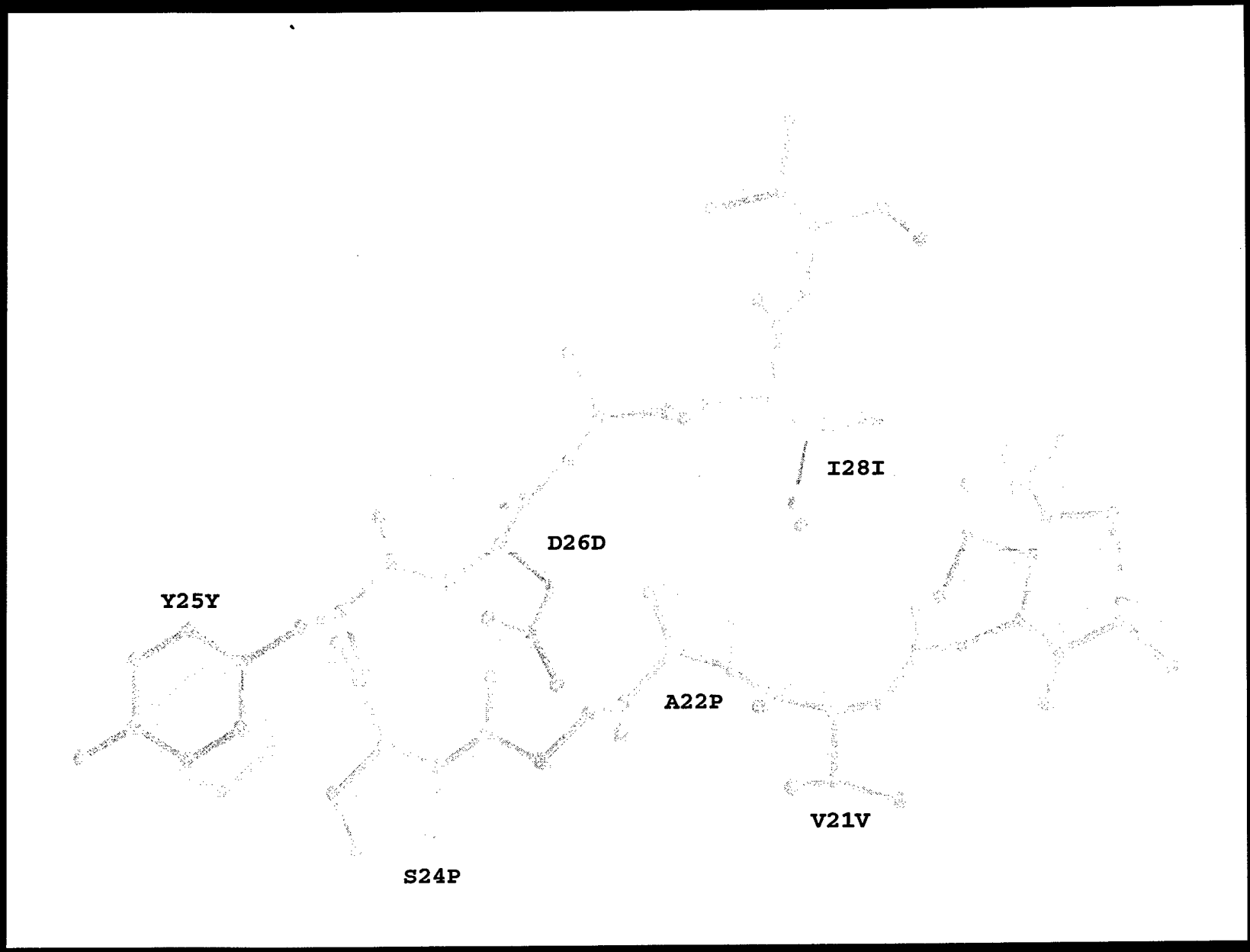

Fig. 5

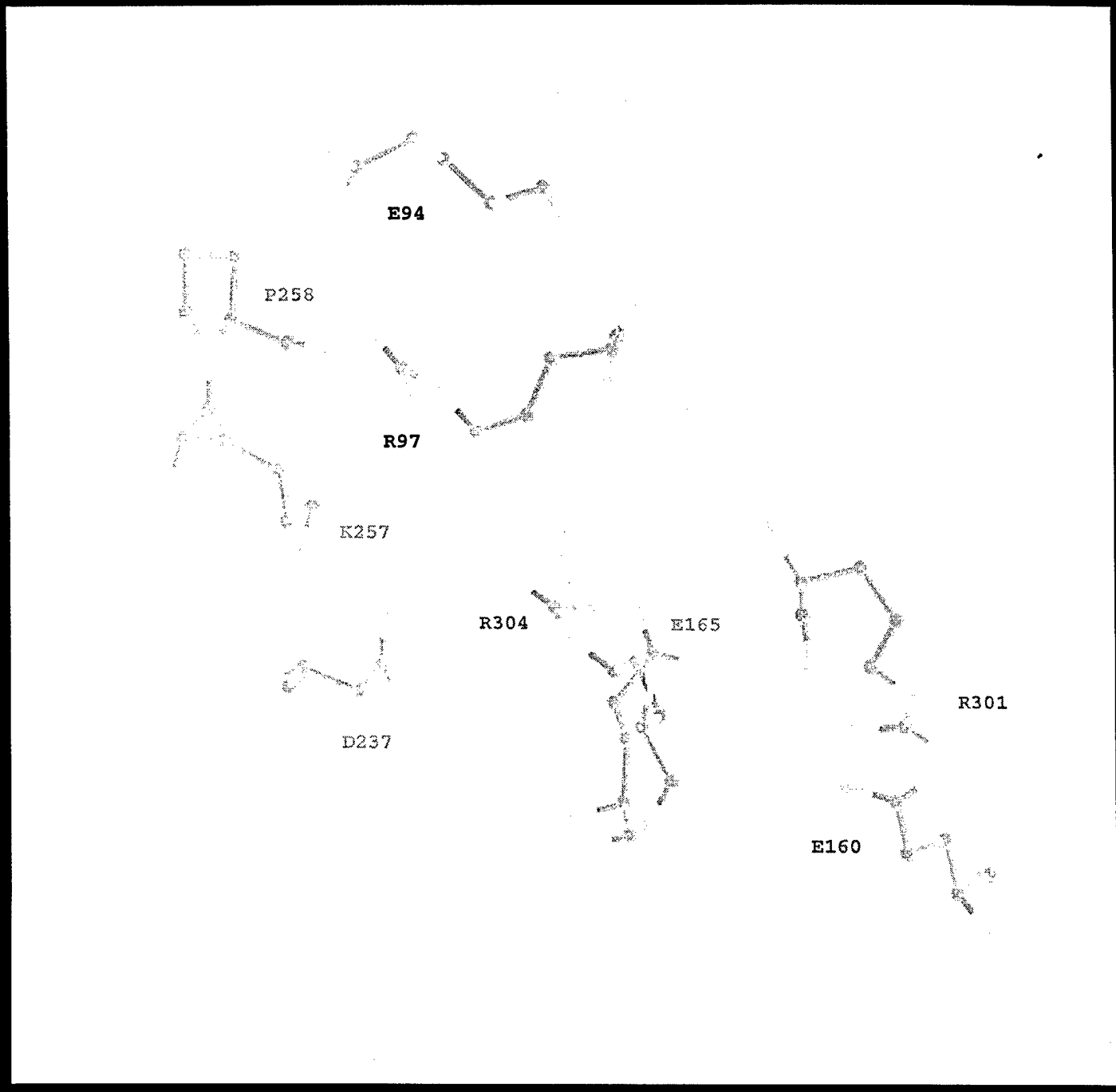

Fig. 6

LIST OF ALL PUBLICATIONS AND TECHNICAL REPORTS

Zhang, Z., Djebli, A., Shoham, M., Frolow, F., Peretz, M., Burstein Y., Crystal parameters of an alcohol dehydrogenase from the extreme thermophile *Thermoanaerobium brockii. J. Mol. Biol.* **230**,353-355 (1993).

Shoham, M., Protein Adaptation to Extreme Envionments. Army Research Office Biosciences Workshop, Cashiers, North Carolina, 12-14 May (1997).

Shoham, M., Protein Adaptation to Extreme Environments. Yellowstone National Park 125th Anniversary Symposium on Life in Extreme Environments. Montana State University, Bozeman, Montana, 16-18 May (1998).

LI, C., Sandriyana, S., H., Joel, and M., Shoham, Crystal structure of a thermophilic alcohol dehydrogenase substrate complex from *Thermoanaerobacter brockii* and comparison to a mesophilic counterpart Submitted for publication

LIST OF ALL PARTICIPATING SCIENTIFIC PERSONNELS

Menachem Shoham, P. I., Associate Professor of Biochemistry

Chunmin Li, Postdoctoral fellow

Sandriyana Soelaiman, M. S. student

Joel Heatwole, undergraduate student